CLAIMS

We claim:

1. A method of reducing the complexity of a first nucleic acid sample to produce a second nucleic acid sample comprising:

fragmenting a first nucleic acid sample to produce fragments;

ligating one or more adaptors to the fragments; and

amplifying a plurality of the fragments by a polymerase chain reaction (PCR), and

modulating the size of the amplified fragments by varying one or more reaction conditions or reagents to reduce the complexity of the first nucleic acid sample.

2. The method of claim 1 wherein the reaction condition or reagent varied is chosen from the group consisting of: extension time, annealing time, primer concentration, primer length, presence or absence of 3' to 5' exonuclease activity, and concentration of nucleotide analogues.

3. The method of claim 1 wherein said adaptors are designed so that the 5' and 3' ends of said fragments are complementary to one another.

4. The method of claim 3 wherein said complementarity is at least 10 bases long and is within 50 bases of the ends of said fragments.

5. The method of claim 3 wherein said complementarity is at least 10 bases long and is within 100 bases of the ends of said fragments.

6. The method of claim 1 where said reaction condition varied is the extension time of said PCR.

7. The method of claim 6 where said extension time is 2-5 seconds.

8. The method of claim 6 where said extension time is 5-10 seconds.

9. The method of claim 6 where said extension time is 10-30 seconds.

10. The method of claim 1 where said reaction condition varied is primer concentration.

11. The method of claim 10 where said primer concentration is 0.1-1 μM.

12. The method of claim 10 where said primer concentration is 0.1-10 μM.

13. The method of claim 10 where said primer concentration is 0.5 to 2.0 μM.

14. The method of claim 10 where said primer concentration is 0.3 to 0.5 μM.

15. The method of claim 1 where said reaction condition varied is primer length.

16. The method of claim 15 where said primer length is 10 to 100 bases.

17. The method of claim 15 where said primer length is 15 to 50 bases.

18. The method of claim 15 where said primer length is 20 to 35 bases.

19. The method of claim 1 where said reaction condition varied is the presence or absence of a 3' to 5' exonuclease activity.

20. The method of claim 1 where said reaction condition varied is the inclusion of one or more strand terminating nucleotides.

21. The method of claim 20 wherein said strand terminating nucleotides are selected from the following dideoxyribonucleotide triphosphates: ddATP, ddTTP, ddGTP, ddCTP and ddUTP.

22. The method of claim 21 wherein the ratio of dNTP to ddNTP is 100 to 1.

23. The method of claim 21 wherein the ratio of dNTP to ddNTP is 1000 to 1.

24. The method of claim 1 further comprising the step of fractionating said fragments according to size by gel filtration chromatography prior to amplification.

25. The method of claim 1 wherein the step of fragmenting said first nucleic acid sample comprises digestion with at least one restriction enzyme.

26. The method of claim 1 wherein the step of fragmenting said first nucleic acid sample comprises digestion with a restriction enzyme that has a six base recognition sequence.

27. The method of claim 1 wherein said adaptor sequences comprise PCR primer template sequences.

28. The method of claim 1 wherein said second nucleic acid sample comprises at least 0.01% of said first nucleic acid sample.

29. The method of claim 1 wherein said second nucleic acid sample comprises at least 0.5% of said first nucleic acid sample.

30. The method of claim 1 wherein said second nucleic acid sample comprises at least 3 % of said first nucleic acid sample.

31. The method of claim 1 wherein said second nucleic acid sample comprises at least 12% of said first nucleic acid sample.

32. The method of claim 1 wherein said second nucleic acid sample comprises at least 50% of said first nucleic acid sample.

33. The method of claim 1 wherein said first nucleic acid sample is genomic DNA, DNA, cDNA derived from RNA or mRNA.

34. The method of claim 1 wherein said size range of a substantial amount of the amplified DNA is 100 to 1000 base pairs.

35. The method of claim 1 wherein said size range of a substantial amount of the amplified DNA is 200 to 1200 base pairs.

36. The method of claim 1 wherein said size range of a substantial amount of the amplified DNA is 400 to 800 base pairs.

37. A method for analyzing a first nucleic acid sample comprising:

obtaining a second nucleic acid sample by:

fragmenting said first nucleic acid sample to produce fragments;

ligating one or more adaptors to said fragments; and,

amplifying the fragments by a polymerase chain reaction (PCR), wherein fragments of a specific size range are preferentially amplified by varying one or more of the reaction conditions or reagents chosen from the group consisting of: extension time, annealing time, primer concentration, primer length, presence or absence of a 3' to 5' exonuclease activity, and concentration of nucleotide analogues;

providing a nucleic acid array;

hybridizing said second nucleic acid sample to said array; and

analyzing a hybridization pattern resulting from said hybridization.

38. The method of claim 37 wherein said method for analyzing a first nucleic acid sample comprises determining whether the first nucleic acid sample contains sequence variations.

39. The method of claim 38 wherein said sequence variations are single nucleotide polymorphisms (SNPs).

40. The method of claim 37 wherein the nucleic acid array is designed to query DNA fragments which have been produced by the procedures used to obtain said second nucleic acid sample.

41. The method of claim 37 wherein a substantial amount of the sequences predicted to be contained in said second nucleic acid sample are predetermined.

42. The method of claim 37 wherein a substantial amount of the sequences predicted to be contained in said second nucleic acid sample are first determined by a computer system.

43. A method of screening for DNA sequence variations in an individual comprising:

providing a first nucleic acid sample from said individual;

obtaining a second nucleic acid sample by:

fragmenting said first nucleic acid sample to produce fragments;

ligating adaptor sequences to said fragments; and

amplifying a subset of the fragments by a polymerase chain reaction (PCR) wherein one or more reaction conditions or reagents are varied to favor amplification of a subset of fragments of a specific size range;

providing a nucleic acid array wherein said array comprises probes designed to interrogate for DNA sequence variations;

hybridizing said second nucleic acid sample to said array;

generating a hybridization pattern resulting from said hybridization; and

determining the presence or absence of DNA sequence variations in the individual based upon an analysis of the hybridization pattern.

44. The method of claim 43 wherein said sequence variation is a single nucleotide polymorphism (SNP).

45. The method of claim 43 wherein the SNP is associated with a disease.

46. The method of claim 43 wherein said SNP is associated with the efficacy of a drug.

47. A method for screening for DNA sequence variations in a population of individuals comprising:

providing a first nucleic acid sample from each of said individuals;

providing a second nucleic acid sample by:

fragmenting said first nucleic acid sample to produce fragments;

ligating adaptor sequences to said fragments; and

amplifying a subset of the fragments by a polymerase chain reaction (PCR) wherein one or more reaction conditions or reagents are varied to favor amplification of a subset of fragments of a specific size range;

providing a plurality of nucleic acid arrays wherein said arrays comprise probes designed to interrogate for DNA sequence variations;

hybridizing each of said second nucleic acid samples to one of said plurality of arrays;

generating a plurality of hybridization patterns resulting from said hybridizations; and

analyzing the hybridization patterns to determine the presence or absence of sequence variation in the population of individuals.

48. The method of claim 47 wherein said sequence variation is a single nucleotide polymorphism (SNP).

49. The method of claim 1 further comprising the steps of diluting the product of said PCR and amplifying the diluted product by a second round of PCR.

50. A method of reducing the complexity of a first nucleic acid sample to produce a second nucleic acid sample whereby said second nucleic acid sample is obtainable by:

fragmenting said first nucleic acid sample to produce fragments;

ligating adaptor sequences to both ends of said fragments such that the 5' and 3' ends of the fragments are complementary to one another; and

amplifying a subset of the fragments by a polymerase chain reaction (PCR) wherein a subset of fragments of a specific size range are preferentially amplified by varying the PCR primer concentration.

51. The method of claim 50 where said primer concentration is 0.1-1 μM.

52. The method of claim 50 where said primer concentration is 0.1-10 μM.

53. The method of claim 50 where said primer concentration is 0.5 to 2.0 μM.

54. The method of claim 50 where said primer concentration is 0.3 to 0.5 μM.

55. A method of genotyping an individual comprising:

identifying a collection of SNPs that are found on fragments of a selected size range resulting from digestion with one or more selected restriction enzymes;

designing an array to interrogate said collection of SNPs;

providing a first nucleic acid sample from said individual;

fragmenting said first nucleic acid sample with the said one or more selected restriction enzymes;

amplifying said fragments by PCR wherein a subset of fragments of said selected size range are preferentially amplified;

hybridizing the PCR product to an array; and

analyzing the hybridization pattern to determine the presence or absence of the collection of SNPs.

56. A kit for reducing the complexity of a nucleic acid sample comprising:

buffer, nucleotide triphosphates, a reverse transcriptase, a nuclease, one or more restriction enzymes, one or more adaptors, a ligase, a DNA polymerase, one or more primers and instructions for the use of the kit.